

GROUND RULES:

- **Print** your name clearly at the top of this page.
- This is a closed-book and closed-notes exam. You can not use external notes of any kind. You may use a calculator.
- This exam contains **two parts**:
 - Part 1. Multiple Choice. 20 questions, 1 point each (20 points total)
 - Part 2. Short Answer. 5 questions, 6 points each (30 points total).

This exam is worth 50 points.

- Any discussion or inappropriate communication between you and another examinee, as well as the appearance of any unnecessary material, will result in a very bad outcome for you (it will be very bad).
- You have **75 minutes** to complete this exam.

HONOR PLEDGE FOR THIS EXAM:

After you have finished the exam, please read the following statement and sign your name below it.

I promise that I did not discuss any aspect of this exam with anyone other than the instructor, that I neither gave nor received any unauthorized assistance on this exam, and that the work presented herein is entirely my own.

MULTIPLE CHOICE. Circle the best answer. Make sure your answer is clearly marked. Ambiguous responses will be marked wrong.

1. Two competing polling organizations, Rasmussen and Gallup, are trying to estimate the proportion of American adults who believe in global warming. Both organizations asked the same question and during the same time frame. Both used a simple random sampling (SRS) design.

The only difference was the sample size used. Rasmussen secured 400 respondents. Gallup secured 2000 respondents. Which organization has a **larger** margin of error associated with its estimate?

- (a) Gallup
- (b) Rasmussen

2. An employee at USC's Student Success Center (SSC) wants to know what percentage of undergraduate students are satisfied with the services offered at the SSC. She takes a **simple random sample** of 100 USC undergraduates and finds that 77 of them are satisfied. In statistics terminology, the number 77% is a

- (a) parameter.
- (b) statistic.

3. A veterinarian is studying new ways to treat mast cell tumors in dogs. She assigns 10 dogs to a standard radiation therapy and 10 other dogs to receive radiation and an experimental drug cocktail. At the end of the study, she will use ultrasound to record how much each dog's tumor has reduced in size.

What are the **individuals** in this study?

- (a) the veterinarian
- (b) the reduction in tumor size
- (c) the dogs
- (d) the form of treatment given (radiation or radiation + drug cocktail)

4. Refer to Question 3. Does this situation describe an experiment or an observational study?

- (a) experiment
- (b) observational study

5. An editorial reporter for *The State* informs his editor that he has received 200 letters from readers on a recent bill making it easier to create charter schools in South Carolina. Among the 200 letter writers, 180 of them oppose the bill. What type of a **sample** is this?

- (a) simple random sample
- (b) cluster sample
- (c) voluntary response sample
- (d) convenience sample

6. The Trump administration has quoted a study that produced an **unbiased estimate** of the proportion of American residents who would like to build a border wall between the United States and Mexico. What does this mean?

- (a) The estimate has a large margin of error.
- (b) The estimate does not have a tendency to overstate or understate the value of the population proportion.
- (c) The estimate has a small amount of uncertainty.
- (d) All of the above.

7. In an experiment, why is it important to use **randomization** when assigning individuals to treatments?

- (a) We would like treatment groups to contain individuals that are similar on average.
- (b) We would like to eliminate confounding variables.
- (c) We would like to minimize the placebo effect for the control group.
- (d) We would like to guarantee our experiment will produce statistically significant results.

8. I take a simple random sample (SRS) of $n = 5$ students from this class. What is the **defining characteristic** of this sample?

- (a) Each sample of 5 students has the same chance of being selected.
- (b) Different students have different probabilities of being selected.
- (c) A student can not appear more than once in the sample.
- (d) The 5 students chosen are guaranteed to be representative of the class.

9. One theory as to why the 2016 presidential pre-election polls were so biased in favor of Hillary Clinton was that certain individuals, when asked who they intended to vote for, would not answer the questions honestly (or they refused to answer them altogether). This is an example of what?

- (a) a sampling error
- (b) a nonsampling error

10. True or False. The **margin of error** for a sample survey includes only random sampling error. It does not include nonsampling errors.

- (a) True
- (b) False

11. Is drinking coffee more efficient in lowering blood alcohol content (BAC) than simply waiting after alcohol consumption? Suppose we recruit 30 subjects to participate in an experiment (18 men and 12 women). Each subject agrees to drink alcohol until their BAC reaches 0.08. At that point, each subject will be randomized to one of the following groups:

- **Group 1:** Drink coffee and wait
- **Group 2:** Do not drink coffee and wait.

Each subject's BAC will then be measured 2 hours later. What is the **response variable**?

- (a) the gender of the subjects
- (b) whether coffee is consumed
- (c) the BAC after two hours of waiting
- (d) the time it takes for a subject's BAC to reach 0.08

12. Refer to Question 11. If you assigned all men to Group 1 and all women to Group 2, what would happen when you analyzed the results?

- (a) You would be able to conclude that drinking coffee reduces BAC.
- (b) You would not be able to distinguish the effect of coffee on BAC from the effect of gender.
- (c) You would get statistically significant results, but the margin of error for the two groups would be large.
- (d) You would not be able to assess the impact of replication.

13. What general statement regarding **statistical inference** is correct?

- (a) We use sample statistics to estimate population parameters.
- (b) We use population parameters to estimate sample statistics.
- (c) We use population statistics to estimate sample parameters.
- (d) We use sample parameters to estimate population statistics.

14. In the language of experiments, what is a **treatment**?

- (a) an experimental condition given or applied to the individuals
- (b) a type of randomization that makes different groups as balanced as possible
- (c) a device used to separate subjects into strata
- (d) a form of statistical control

15. The last time I went to Atlanta, I decided to catch a Braves game. As I looked throughout the stands, I wondered what proportion of attendees were Braves fans. To answer this question, I asked 10 people in the row in front of me if they were Braves fans. What type of **sampling design** did I use?

- (a) a stratified sample
- (b) a simple random sample
- (c) a cluster sample
- (d) a convenience sample

16. A simple random sample of $n = 400$ Birmingham, AL residents was recently asked

Have you ever been the victim of a hate crime?

From this sample, 100 residents responded “Yes.” What is the value of the **sample proportion** of hate-crime victims?

- (a) 0.05
- (b) 0.10
- (c) 0.20
- (d) 0.25

17. What type of sample will produce a zero margin of error?

- (a) a simple random sample
- (b) a census
- (c) a cluster sample with secondary-stage stratification
- (d) a non-probability sample

18. The Ministry of Health in the Canadian province of Ontario wants to know whether citizens of Ontario approve of their health care services. A statistician selects the three largest cities in Ontario (Toronto, Ottawa, and Hamilton). From each city, he selects a simple random sample (SRS) of 100 residents who have visited the emergency room in the last year.

After reporting the results for his sample to his boss, the statistician is fired. Why do you think he was fired?

- (a) A sample of size 300 is not nearly large enough to learn about all of Ontario's residents.
- (b) The response variable in the survey is not quantitative.
- (c) The sample he took suffers from undercoverage.
- (d) The margin of error associated with a sample size of 300 is too small to produce statistically significant results.

19. A medical researcher obtained blood specimens from a simple random sample (SRS) of 31 young children in New Delhi, India, all of whom were infected with malaria. What group of individuals is the most reasonable choice to serve as the **population** in this setting?

- (a) all children in New Delhi, India, who have malaria
- (b) all children in India
- (c) the 31 children selected
- (d) all children in New Delhi, India

20. You want to conduct a survey involving undergraduate students at USC. You would like to guarantee equal representation of freshmen, sophomores, juniors, and seniors. Which **sampling design** should you use?

- (a) a simple random sample
- (b) a systematic sample
- (c) a stratified random sample
- (d) a voluntary response sample

SHORT ANSWER. Give detailed responses. Please write clearly and legibly.

1. The Denver Police Department wants to know if Hispanic residents of Denver believe that the police use racial profiling when making traffic stops. A sociologist prepares several questions about the police. The police department chooses a simple random sample (SRS) of 200 mailing addresses in “predominantly Hispanic neighborhoods” and sends a uniformed Hispanic police officer to each address to ask the questions of an adult living there.

- (a) What is the population? What is the sample?
- (b) Give one source of sampling error.
- (c) Identify two nonsampling errors that could arise.

2. During January 23-24, 2017, Rasmussen Reports conducted a national telephone and online survey using a simple random sample (SRS) of $n = 1000$ American adults. Each participant was asked:

Will the recent Women's March on Washington be good for women around the world?

The survey found that 420 of the 1000 adults in the sample answered "Yes" to this question. Use this information to write a **95 percent confidence statement**.

Note: A confidence statement is a well-written sentence. Show all of your calculations.

3. I asked you to read Judith Singer's paper, "Afraid to Discuss Evolution?," and we discussed this paper in class. The paper summarized a 2005 *New York Times* editorial. Summarize the findings of Singer's paper and our classroom discussion. In particular,

- state the topic discussed in the *New York Times* editorial
- state at least three things that led to Singer's attack of the editorial board's conclusions
- state at least one thing you learned from our discussion (e.g., about the media, etc.).

Successfully addressing each of these topics will earn you full credit. Use clearly written and complete sentences.

4. McCann and Tebbs (2009) summarize a study examining perceived unmet need for dental health care for people with HIV infection. Baseline in-person interviews were conducted with 2,864 HIV infected individuals (aged 18 years and older) as part of the HIV Cost and Services Utilization Study.

(a) One of the questions asked was

“In the last six months, was there a time when you needed dental treatment but could not get it?”

Assuming that the 2,864 individuals form a simple random sample, calculate the margin of error associated with the sample proportion who respond “Yes” to this question. Assume a 95% confidence level.

(b) Identify two stratifying variables that may be of interest to the investigators in this study. Explain why.

(c) Should the question above have been asked using the randomized-response technique? If so, explain why. If not, say why not.

5. Omeprazole is used to treat symptoms of gastroesophageal reflux disease and other conditions caused by excess stomach acid. A chemist has developed a new drug that she believes does a better job at treating these same symptoms and conditions.

The chemist would like to design a randomized comparative experiment that has three treatment groups:

- **Group 1:** patients who will take a standard dose of omeprazole
- **Group 2:** patients who will take a standard dose of the new drug
- **Group 3:** patients who will take a placebo.

She has 90 patients who will participate and she would like to have 30 patients per treatment group.

- (a) How should the patients be assigned to the different groups? Explain precisely how you would do this.
- (b) List one potential lurking variable that could arise.
- (c) What should the response variable be?
- (d) If the chemist found the results of this experiment to be statistically significant, what does this mean in practical terms?